file copy

A Molecular Portrait of Human Papillomavirus Carcinogenesis

T.R. Broker,* L.T. Chow,* M.T. Chin,* C.R. Rhodes,* S.M. Wolinsky,[†] A.Whitbeck,[‡] and M.H. Stoler[‡]

*Biochemistry Department, Cancer Center; [†]Department of Medicine, Infectious Diseases Unit; [‡]Pathology Department, Cancer Center; University of Rochester School of Medicine, Rochester, New York 14642

The human papillomaviruses (HPVs) specify a complex family of overlapping polycistronic mRNAs, including one that encodes transformation proteins and an autoregulatory transcriptional factor capable of both positive and negative feedback. The viral enhancer-promoter region also responds to constitutive and conditional host transcriptional factors. Together these regulatory proteins modulate viral transcription in coordination with epithelial tissue differentiation. We have designed probes that can distinguish each of the five prevalent and related HPVs that infect the anogenital region. We have also built subgenomic clones from which riboprobes specific for individual mRNA species are generated in vitro. These probes reveal the viral transcription patterns in serial sections of patient biopsies after in situ hybridization. Our studies on a spectrum of lesions ranging from benign epithelial hyperproliferation to invasive carcinomas associated with HPV types 6, 11, 16, and 18 demonstrate that viral gene expression is tightly linked to cellular differentiation. The most definitive parameter in characterizing the state of viral-host interactions associated with the oncogenic HPV types 16 and 18 resides in the relative abundances of mRNAs from the E6-E7, E4-E5, and L1 regions. In low-grade lesions, all are expressed, with the E4-E5 RNA being most abundant. As the severity of neoplasia increases, the expression of E4-E5 and L1 decreases or becomes absent. Conversely, the E6-E7 region is derepressed. On the basis of this information, we propose a molecular mechanism for the derepression of the viral transformation genes. The recently described interactions between the viral transformation protein and the retinoblastoma (RB) anti-oncoprotein and perhaps others probably also play key roles in the initiation of viral carcinogenesis (Dyson et al. 1989 and this volume).

Neoplasia of the uterine cervix has traditionally been classified according to histopathologic criteria (Patten 1978; Koss 1979). Intraepithelial neoplasms appear to form a morphologic continuum, with clear evidence for progression through successively more severe stages to invasive cancer. Virtually all of these morphologic alterations have been consistently associated with one or more types of HPV infection. Nearly two dozen HPV types have been found in the anogenital tract. HPV types 6 and 11 are most often associated with benign venereal warts (condylomata acuminata) in the lower genital tract and are also infrequently found in some carcinomas. In contrast, higher-grade intraepithelial neoplasia, carcinomas in situ, and invasive carcinomas are associated primarily with HPV types 16, 18, 31, 33, and 35 (for review, see Pfister 1987). This paper describes molecular studies of HPV transcription and regulation, relates them to the transcription profiles observed in a spectrum of HPV-16- and HPV-18-associated genital tract lesions and proposes a molecular mechanism for HPV-induced carcinogenesis.

[†]Present address: Northwestern University School of Medicine, Department of Medicine, Infectious Diseases Unit, Chicago, Illinois 60611.

The normal epithelium of the uterine cervix varies according to anatomic site. The lining of the endocervical canal is a mucus-secreting, unstratified columnar epithelium. The ectocervix is covered by a stratified noncornified squamous epithelium. After puberty, the position of the squamo-columnar junction generally moves up the endocervical canal as glandular epithelium becomes replaced by metaplastic squamous epithelium. This transformation zone, where active epithelial turnover is most evident, is also the site for initiation of most of the epithelial neoplasms of the cervix. All of the epithelial cell types are probably derived from a common pool of dividing multipotential reserve cells that can differentiate along a variety of paths, depending on regional stimuli (Coppleson and Reid 1967; Gould et al. 1979). In squamous mucosae, the dividing cell population is generally restricted to the basal and possibly parabasal layers. A subpopulation of daughter cells are pushed up toward the surface. These cells undergo several successive changes in the expression of particular subsets of the keratin gene family (Moll et al. 1983). Such differentiation is reflected in the characteristic cellular morphology in a given epithelial cell layer. Under normal conditions in the ectocervix, the mature squamous mucosa does not have a well-defined granular or keratin

CANCER CELLS 7 / *Molecular Diagnostics of Human Cancer.* 0-87969-324-X/89 $1.00

layer. The superficial squamous cells contain abundant cytoplasm and a small pyknotic nucleus and are the primary cells sloughed to the environment. In contrast, external cutaneous skin matures several steps further to produce a highly cornified surface layer.

Papillomaviruses induce changes in differentiating epithelia. The most pathognomonic cell, often seen in productive infections, is called a koilocyte (koilo means cave), which has a large vacuole around the nucleus. The nuclei of koilocytes are almost always atypical in that they are enlarged and hyperchromatic, with an associated increase in nuclear to cytoplasmic ratio. HPVs also induce architectural changes in the epithelium, including papillary hyperproliferation and epithelial thickening (acanthosis), thought to result from the enlargement of individual cells and from a delayed course of differentiation leading to the accumulation of excessive numbers of cells prior to desquamation (Steinberg 1986). These features are characteristic of the usually benign exophytic neoplasm, condyloma acuminatum of the vulva and vagina. However, in the cervix similar changes are more frequently associated with flat epithelial proliferations lacking papillary growth. Such "flat condylomas" are recognized as the earliest manifestation of a spectrum of cervical dysplasia (intraepithelial neoplasia). Morphologic progression through the stages of dysplasia includes replacement of increasing proportions of the epithelium above the parabasal stratum with dividing cells that often exhibit nuclear abnormalities, such as increased numbers of mitoses, abnormal mitoses (aneuploidy), and multinucleated cells. There is concurrently a proportional decrease in the frequency of koilocytes, presumably because of the absence of the differentiation signals necessary for the production of viral cytopathic effects associated with late viral gene transcription. Dysplasias are graded as mild (cervical intraepithelial neoplasia grade I [CIN I]), moderate (CIN II), or severe (CIN III) corresponding, respectively, to the replacement of one-third, less than two-thirds, or greater than two-thirds of the epithelium by basal-like cells. Classic carcinoma in situ of the cervix (also called CIN III) is recognized as a stage in which the entire epithelial thickness is replaced by primitive undifferentiated cells.

Squamous carcinomas of the cervix generally arise from, or in association with, squamous intraepithelial neoplasms. They may have keratinizing, nonkeratinizing, or undifferentiated morphology, and may be composed of large or small cells exhibiting a range of individual cell differentiation. Adenocarcinomas of the cervix arise in glandular epithelium, often associated with a morphologic precursor, in situ adenocarcinoma. Small-cell undifferentiated carcinomas do not have a well-defined histogenetic precursor and frequently demonstrate neuroendocrine differentiation. They are analogous to small-cell carcinomas of the lung and are characterized by a particularly aggressive clinical course. The developmental relationships among these three types of epithelial differentiation are not certain, but it appears that the reserve cell population is capable of differentiating along all three directions. Initiation of neoplastic development could conceivably occur in a cell already committed to one of the differentiation paths. Alternatively, carcinogenesis may be initiated in a common multipotent precursor cell, with the subsequent differentiation influenced by genetic and environmental events during tumor progression. Because tumors with mixed differentiation are not unusual, a common stem cell capable of multipotential differentiation is quite possible. We present some additional evidence based on patterns of HPV gene expression in favor of a common origin of these tumors. Similar patterns of histopathogenesis are recapitulated in other anatomical sites, including the male urogenital tract, the anal mucosa, and the nasal, laryngeal and respiratory mucosa, all of which have morphologically similar transformation zones and are subject to HPV infection.

Papillomavirus genome structure, transcription, and regulation

Numerous human and animal papillomavirus genomes have been cloned, and for some, their DNA sequences have been determined. They are a family of related double-stranded DNA viruses with circular genomes of approximately 7900 bp that replicate as extrachromosomal plasmids in the nuclei of epithelial cells in benign lesions. Their genomes are all similarly organized into early and late open reading frames (ORFs) encoded by the same DNA strand (Fig. 1C). Immediately preceding the early ORFs is an upstream regulatory region (URR) or a long control region, which contains transcriptional enhancer elements, promoters, and DNA replication control sequences (for review, see Broker and Botchan 1986). The early (E) region encodes *trans*-acting factors required for regulated, extrachromosomal replication (E1 ORF), enhancer activation and repression (E2 ORF), and cellular transformation (E5, E6, and E7 ORFs). The viral capsid proteins (L1 and L2) are the products of the late region and are expressed in productive lesions but not in transformed cells.

In benign HPV-infected lesions, the viral DNAs exist as extrachromosomal plasmids, mostly as monomeric circular molecules. However, in some but not all cancers associated with HPV-16, viral DNAs are found as multimeric circular molecules, sometimes with deletions (Dürst et al. 1985; Choo et al. 1987; Smotkin and Wettstein 1987). In other cancers, viral DNA is integrated into host chromosomes (Dürst et al. 1985, 1987; Matsukura et al. 1986; Shirasawa et al. 1986; Choo et al. 1987). Viral integration, when characterized, invariably disrupts the E2 ORF encoding the transcription regulatory proteins (Schwarz et al. 1985; Matsukura et al. 1986; El Awady et al. 1987; Baker et al. 1987; Choo et al. 1988). Naturally occurring malignant lesions express viral E6 and E7 mRNAs (Smotkin and Wettstein 1986; this paper). In addition, most long-established cervical carcinoma cell lines, such as HeLa, SiHa, and CaSki, have been found to harbor integrated HPV types 16 or 18 DNA from which the transforming E6 and E7 regions are actively transcribed (Pater and Pater 1985;

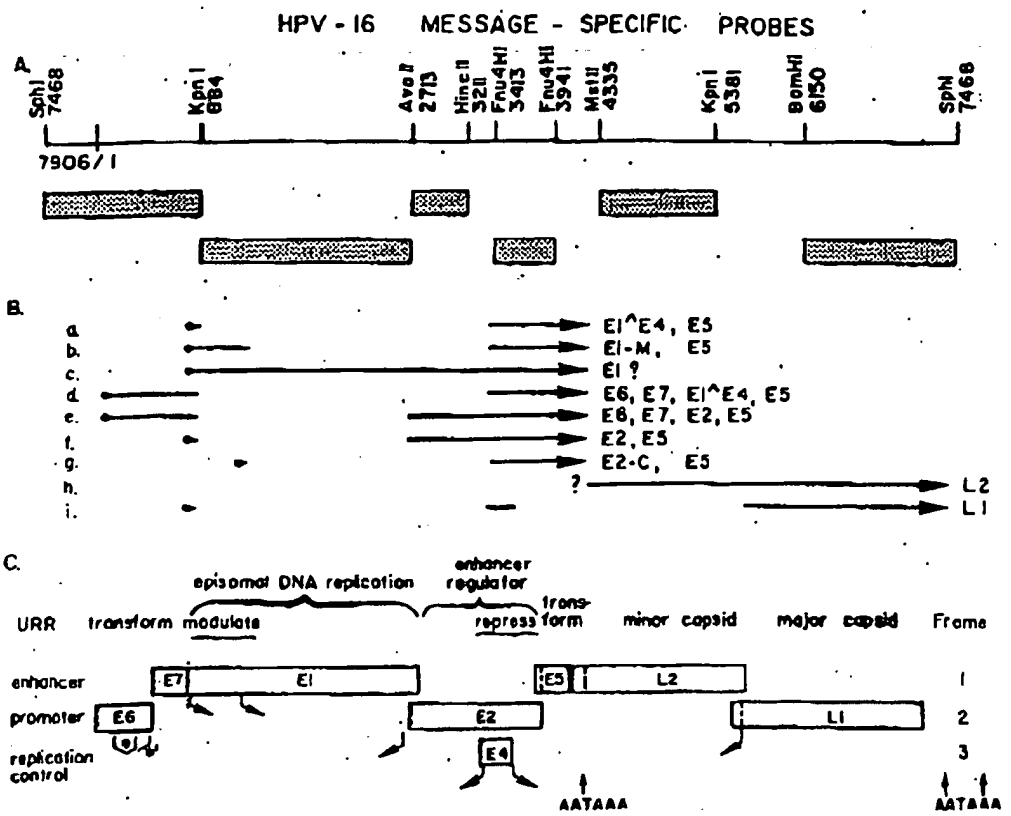

Figure 1 HPV-16 exon-specific probes. (*A*) Subgenomic segments of HPV-16 DNA corresponding to the anticipated mRNA exons were cloned into pGEM dual promoter vectors. (*B*) A representative set of HPV-6 and HPV-11 mRNAs (Chow et al. 1987a) and their probable protein-coding potential. (*C*) The HPV-16 ORFs based on DNA sequence (Seedorf et al. 1985) and the probable functions of the encoded proteins. RNA splice donor and acceptor sites when known or inferred from homology with HPV-11 are marked with bent arrows.

Schwarz et al. 1985; Yee et al. 1985; Schneider-Gädicke and Schwarz 1986; Baker et al. 1987). These observations further strengthen the important relationship between these HPV types and cervical neoplasia.

Cellular transformation in vitro also has pointed to an active role for certain papillomavirus types in these processes. HPV-16 DNA immortalizes cultured primary foreskin keratinocytes or primary cervical cells in culture (Pirisi et al. 1987; Woodworth et al. 1988). HPV types 16, 18, 31, and 33, but not HPV types 6 or 11, are capable of transforming primary baby rat kidney epithelial cells in collaboration with an activated cellular oncogene, Ha-*ras* (Storey et al. 1988), thus mimicking the multistep carcinogenesis experiments described by Land et al. (1983). HPV-16 also inhibits the ability of primary keratinocytes to differentiate when cultured on collagen rafts at the air-media interface (Asselineau and Prunieras 1984; Kopan et al. 1987), inducing morphologic transformation that mimics CIN (McCance et al. 1988). In such experimental systems, the transformed phenotype is not apparent until the cells have been passaged many generations, again suggesting the need for additional genetic events. The major transformation gene of HPV-16 is E7 (Phelps et al. 1988; Storey et al. 1988). The E7 protein has been identified in cervical carcinoma cell lines (Smotkin and Wettstein 1986) and in transformed rodent cells (Banks and Crawford 1988). Recently, a homology between the HPV E7 protein and other viral and cellular oncogenes has been recognized (Phelps et al. 1988). Moreover, these proteins form complexes with the recently characterized RB anti-oncoprotein (Lee et al. 1987; DeCaprio et al. 1988; Whyte et al. 1988; Dyson et al. 1989 and this volume), suggesting the interaction between E7 and RB proteins could play a role in HPV-induced maligant transformation. It is not known whether these observations are related to the elevated c-*myc* and c-*ras* expression in some cervical carcinomas (Riou et al. 1985).

RNA transcription and regulation

We have characterized the structures of HPV types 6 and 11 mRNAs by electron microscopy of RNA:DNA heteroduplexes (Fig. 1B) (Chow et al. 1987a) and those of HPV types 11, 16, and 18 by cDNA analysis (Naeseri et al. 1987; M.O. Rotenberg et al., in prep.; D. Palermo-Dilts et al., unpubl.). The critical polycistronic E2 mRNA on which this study is centered originates from a highly conserved TATA motif at the beginning of the early

region, the E6 promoter, which is characteristic of all HPVs trophic for the genital tract. It spans the E6 and E7 ORFs and, shortly into the E1 ORF, is spliced to a site just upstream from the E2 ORF and continues to the early polyadenylation site just downstream from the E5 ORF (Fig. 1B, e). Although encoded by the third major ORF in the spliced message, the E2 protein is synthesized in various mammalian cell lines (M.O. Rotenberg et al., in prep.). An E2-C mRNA (Fig. 1B, g) derived from a promoter in the middle of the E1 ORF has a short 5′ exon with the initiation codon spliced to a major exon coding for the carboxy-terminal half of the E2 ORF (E2-C). Similar mRNAs for HPV-16 and HPV-18 are predicted based on genomic DNA sequence homology and conservation of splice sites and coding capabilities. Multiple forms of E2 proteins have also been reported for bovine papillomavirus type-1 (Lambert et al. 1987; Hubbert et al. 1988). Both E2 and E2-C proteins regulate the enhancer and E6 promoter (Spalholz et al. 1985; Hirochika et al. 1987; Lambert et al. 1987; Phelps and Howley 1987; Chin et al. 1988 and in prep.). We have dissected the HPV-11 regulatory sequences in the URR and have shown that it consists of three components, a set of sequences responsive to E2 or E2-C proteins (E2-RS), a constitutive enhancer I (CEI), which has no cell-type specificity, and a constitutive enhancer II (CEII), which functions only in epithelial cells of human origin (Fig. 2) (Hirochika et al. 1987; Chin et al. 1988 and in prep.). We and other investigators (Androphy et al. 1987; Giri and Yaniv 1988; McBride et al. 1988; Moskaluk and Bastia 1988) have shown that the DNA-binding domains of E2 and E2-C proteins of human and animal papillomaviruses are in the common carboxy-terminal portions. Each recognizes the E2-RS with a sequence motif $ACCN_6GGT$, which occurs several times in the URR of all papillomavirus types, both human and animal, accounting for enhancer regulation by heterologous E2 proteins (Hirochika et al. 1987). In the mucosotrophic HPVs, two tandem copies (Fig. 2, nos. 3 and 4) of the E2-responsive sequence immediately precede the E6 promoter TATA motif, and a third copy (Fig. 2, no. 2) is further upstream near CEI. Purified HPV-11 E2 and E2-C proteins expressed in *Escherichia coli* can bind to each copy, and their DNAse I footprints extend beyond the core motif (Chin et al. 1988; Hirochika et al. 1988). When the tandem E2-responsive sequences (Fig. 2, nos. 3 and 4) are occupied in the absence of CEII, E2 protein acts as a transcriptional repressor rather than as an activator (M.T. Chin et al., in prep.), presumably because most of the TATA motif is occluded and unable to bind host transcription factor TFIID (Sawadogo and Roeder 1985). If the TATA motif is located further away from E2 responsive sequences, as it is in the human and animal papillomaviruses trophic for cutaneous skin or in recombinant plasmids containing the HPV URR linked to the SV40 promoter, the full-length E2 protein acts as a transcriptional *trans*-activator (Thierry and Yaniv 1987; Chin et al. 1988). The unique amino-terminal portion of each of the full-length E2 proteins has a highly conserved amphipathic, acidic α-helical domain analogous to those of many prokaryotic and eukaryotic transcription factors and is believed essential for protein-protein interactions in the formation of active transcription complexes. E2-C proteins invariably act as repressors, presumably by competing for E2-responsive sites without bringing in the activating domain (Cripe et al. 1987; Lambert et al. 1987; Chin et al. 1988).

The incorporation of tandem copies of synthetic CEII sequences 38-bp long into a URR deletion mutation lacking the native CEI and CEII partially abrogates E2 repression of the E6 promoter (M.T. Chin et al., in prep.). Similarly, multiple copies of a restriction fragment containing CEI also overcome the E2 repression and constitute a strong E2-independent enhancer in different cell types (Hirochika et al. 1988; M.T. Chin et al., in prep.). In cervical carcinoma cell lines, we have identified a ~44,000-dalton protein that binds to CEII.

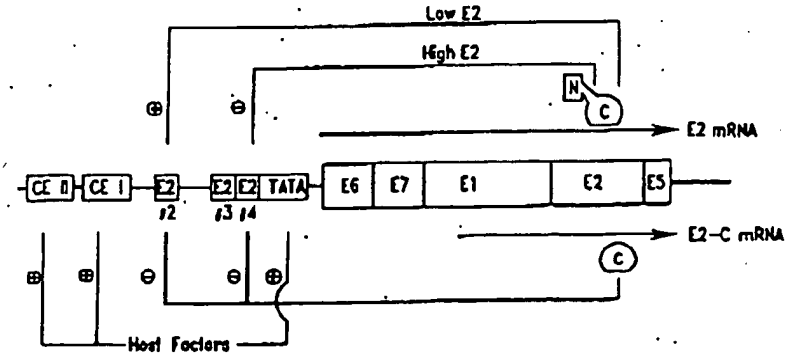

Figure 2 Regulation of early region transcription of HPVs trophic for the genital tract. Most of the features are known from studies of HPV-11 (Chow et al. 1987a; Hirochika et al. 1987, 1988; Chin et al. 1988 and in prep.). Cellular enhancer elements I (constitutive) and II (epithelial cell specific) vary among HPV types. HPV types 6, 11, 16, 18, 31, 33, and 7 all have tandem E2-responsive sequences (copies *3* and *4* for HPV-11, used as a model here) adjacent to the E6 promoter TATA motif and have another (copy *2*) near the cellular enhancer elements. The structures of the polycistronic E6-E7-E2-E5 and the E2-C-E5 mRNAs are shown. Both encode the carboxy-terminal DNA-binding domain of the E2 proteins, but only the full-length E2 transcript specifies the amino-terminal activation domain. The positive and negative effects of various host and viral transcriptional regulatory proteins are indicated by + and −, respectively, with the E2 protein being the key to the feedback circuit. The thin line in each mRNA represents an intron.

Although we have found that purified E2 and E2-C proteins expressed in *E. coli* have essentially equal affinities for the different copies of the E2-RS in the absence of other cellular factors, we speculate that E2 protein may preferentially bind to E2-RS 2 (Fig. 2) in association with these host factors, leading to the activation of the E6 promoter. Upon continued expression of the operon, we propose that E2 protein levels rise to a concentration sufficient to occupy E2-RS 3 and 4 (Fig. 2), blocking the TATA motif and down-regulating its own production to a low, steady state. In doing so, we believe that E6, E7, and possibly E5 transforming protein concentrations are also controlled. We also postulate that this feedback mainly occurs in the less differentiated basal and parabasal cells in which the CEI- and CEII-binding proteins might be limiting, thus preventing premature killing of the infected stem cells. This autoregulation is presumably overcome by increased concentration of these host factors in the differentiated keratinocytes, allowing the virus to proceed into productive infection. Such coordinate positive and negative feedback regulation is a hallmark of the bacteriophage λ repressor gene and of the *E. coli* arabinose operon, for instance, but this would appear to be one of the first candidates in a eukaryotic transcription unit. The balancing role played by E2-C repressor in this regulation is not yet clear. E2 proteins expressed in *E. coli* form dimers, including E2:E2-C heterodimers (McBride et al. 1988), potentially adding another level of fine tuning that can only be evaluated when regulation of the E2-C promoter is understood.

In situ hybridization

Exploration of the association between viral gene expression and cellular differentiation demands the ability to discern topographical differences of transcriptional activity in the context of tissue morphology. Such information is invariably lost by mass tissue analysis. The amount of tissue necessary for such biochemical analysis is relatively large and is further complicated by the number of viral types, viral genes, and host genes that need to be examined. Furthermore, pathological examination cannot be performed on the same tissue that is analyzed biochemically. In situ hybridization is the only technology that can address each of these problems. It permits the sensitive detection of specific mRNA and DNA sequences in recent and archival, formalin-fixed biopsies, while preserving the histology of the tissue and revealing the exact cellular and subcellular location of the sequences under study. Over the past 3 years, these methods have been increasingly utilized to address the molecular association of HPVs with neoplasia, as well as being recognized as one of the most powerful diagnostic and research tools available in anatomic pathology.

HPV cross-reactive and type-specific probes

We have described previously DNA:DNA heteroduplex mapping of genital HPVs and shown that the L1 ORFs are the most highly conserved regions, whereas the E5 region and the URR-E6-E7 segment are most unique to each type (Broker and Chow 1986; Chow et al. 1987b). On the basis of these data, subgenomic probes that can either serve as papillomavirus group cross-reactive probes or as type-specific probes were designed and confirmed by application to Southern transfer blots of DNA prototypes (Chow et al. 1987b; see also Manos et al., this volume). Single-stranded RNA probes generated from these clones were used to examine a wide variety of patient biopsies by in situ hybridization to conventional formalin-fixed, paraffin-embedded serial sections with a high sensitivity of detection. They generally reveal foci of infection by a single HPV type, but occasionally two types may be present, usually coinfecting the same clusters of cells, as though one of the viruses may provide a helper effect for the other.

An example of cross-reaction of RNA probes corresponding to the whole viral genomes of the closely related HPV-16 and HPV-31 (Lorincz et al. 1986; Chow et al. 1987b) is shown in Figure 3. Even at $T_m - 5°C$, HPV-31 whole genomic probes yield a weak positive signal with HPV-16 RNA present in a biopsy of a CIN I. When a subgenomic probe of HPV-16 from the relatively unique URR-E6-E7 region was used, a strong signal was maintained, whereas an HPV-31 E2-C-E5-L2-N type-specific probe did not cross-react. Similar type-specific RNA probes of HPV types 6, 11, 16, 18, and 31 (Chow et al. 1987b; M. Stoler et al., in prep.) have been used to identify and type papillomaviruses in a large number of lesions using in situ hybridization. This paper focuses on lesions containing HPV-16 and HPV-18.

HPV-16 mRNA expression revealed by in situ hybridization with exon-specific probes

There is only limited information on the precise structures of HPV-16 mRNAs, predominantly from S1 nuclease mapping of E6-E7 early region transcripts recovered from CaSki cells (Smotkin and Wettstein 1986), a cervical carcinoma cell line known to contain several hundred copies of full-length HPV-16 DNA integrated in tandem into several host chromosomes. Nonetheless, the similarity of genomic organization of all papillomaviruses and the conservation of splice donors, splice acceptors, and polyadenylation sites evident in the HPV-16 DNA sequence suggest that the mRNA structures are likely to be analogous to transcripts of HPV types 6 and 11 (Chow et al. 1987a). The exception is the presence of small introns in the E6 transcripts of HPV-16 and HPV-18 as well (Schneider-Gädicke and Schwarz 1986; Smotkin and Wettstein 1986). Using polymerase chain reactions on cDNAs prepared from RNA isolated from CaSki cells, we have identified a common splice donor (nt 226) and two alternative splice acceptors (nt 409 and nt 526) in HPV-16 transcripts. Both species lead to truncation of E6 protein by changing into reading phases closed shortly after the splice acceptor site (D. Palermo-Dilts et al., unpubl.).

We have cloned subgenomic regions of HPV-16 corresponding to the anticipated mRNA exons (Fig. 1A). These were placed in dual promoter vectors from which

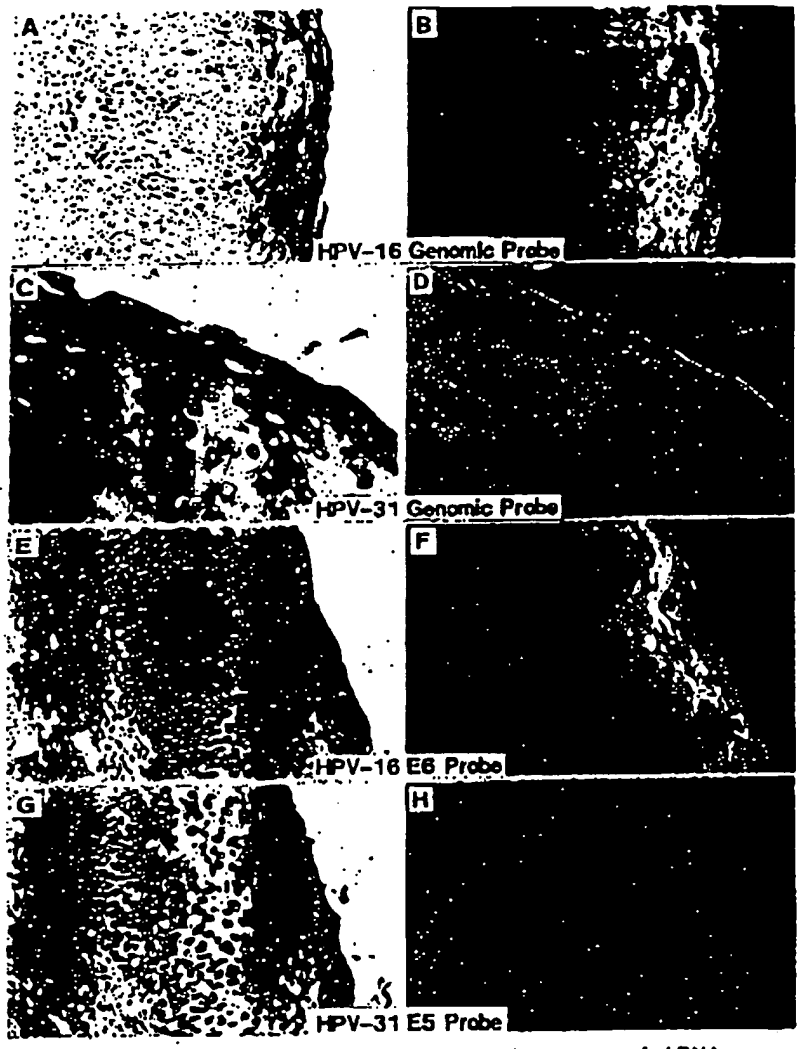

Figure 3 HPV-16 type-specific probes. ^{3}H-labeled antisense riboprobes complementary to viral RNA were generated from full-length genomic HPV-16 or HPV-31 DNA or from a subgenomic HPV-16 URR-E6-E7 (nt 7468–884) or subgenomic HPV-31 fragment spanning the region from E4/E2-C through the amino-terminal portion of the L2 ORF (nt 3362–4584). They were hybridized to adjacent sections of a biopsy diagnosed as CIN I. The sections exhibit HPV RNA signals when either whole genomic probe was employed. In contrast, only the HPV-16 RNA signals are evident when subgenomic, type-specific probes were applied.

either sense-strand or antisense-strand RNA probes can be generated by in vitro transcription. Tritium-labeled sense-strand RNA probes can be hybridized to denatured viral DNA in lesions without risk of confounding hybridization to viral mRNA of the same polarity. Conversely, ^{3}H-labeled antisense RNA will anneal with high efficiency and specificity to viral RNA in samples not subjected to prior denaturation of the DNA. Thus, the DNA and RNA distributions in specimens can be clearly distinguished.

Serial sections of biopsies of various cervical dysplasias and carcinomas were processed for in situ hybridization to probe for viral DNA or RNA, as described previously (Stoler and Broker 1986). Sections were first screened with HPV types 6, 11, 16, 18, and 31 type-specific probes. Hybridization was for 12–16 hours at T_m – 25°C; excess probe was eliminated by digestion with ribonuclease and high-stringency washing at T_m + 5°C. Sections were overlaid with liquid photographic emulsion and autoradiographed for 4 weeks. Following development, the slides were observed and photographed by dark-field illumination, which produces dramatic light scattering from ^{3}H-exposed silver grains and hence increases the sensitivity of detection. Signals are localized to the nucleic acid source: DNA and RNA precursors or processing by-products are in the nuclei, whereas mRNAs are predominantly in the cytoplasm.

Specimens containing HPV-16 that are described here include one case of koilocytotic atypia/CIN I, one case of CIN III, a set of three biopsies from a single

patient (representing squamous carcinoma in situ and endocervical adenocarcinoma in situ from a cone biopsy, as well as invasive adenosquamous carcinoma in the uterine corpus) and a second invasive squamous carcinoma. Together, these represent a nearly complete pathological spectrum. A set of serial sections of each specimen were challenged with mRNA exon-specific probes specific for E6-E7, E1, E2, E4(E2-C)-E5, L2 and L1 ORFs. The amounts of probes used were normalized according to size so that the signal intensities represent relative copy numbers of the target RNA species. Regions of each slide showing the same histological features were photographed and also semiquantitatively evaluated by photodensitometry of light scattering. One section of each specimen was hematoxylin plus eosin stained for histopathological evaluation. A subset of the autoradiographs from this study are presented in Figures 4, 5, and 6.

All early as well as late exons were abundant in the CIN I, with a distribution identical to that typically seen in HPV-6- or HPV-11-associated vulvar condylomas (M. Stoler et al., in prep.). E6-E7 and E4-E5 exon signals first appeared just above the basal cell layer of the epithelium (Fig. 4B,E). There was a dramatic increase in signal intensity in the more differentiated keratinocytes. This is the same transition point where viral DNA copy (replication) also becomes abundant (data not shown). One clear and perhaps significant difference between this HPV-16 CIN I and HPV-6 and HPV-11 vulvar condylomas is that the HPV-16 E4-E5 exon signal is only 2–3 times more abundant than the E6-E7 exon signal, whereas it is 10–20 times more abundant in HPV-6 and

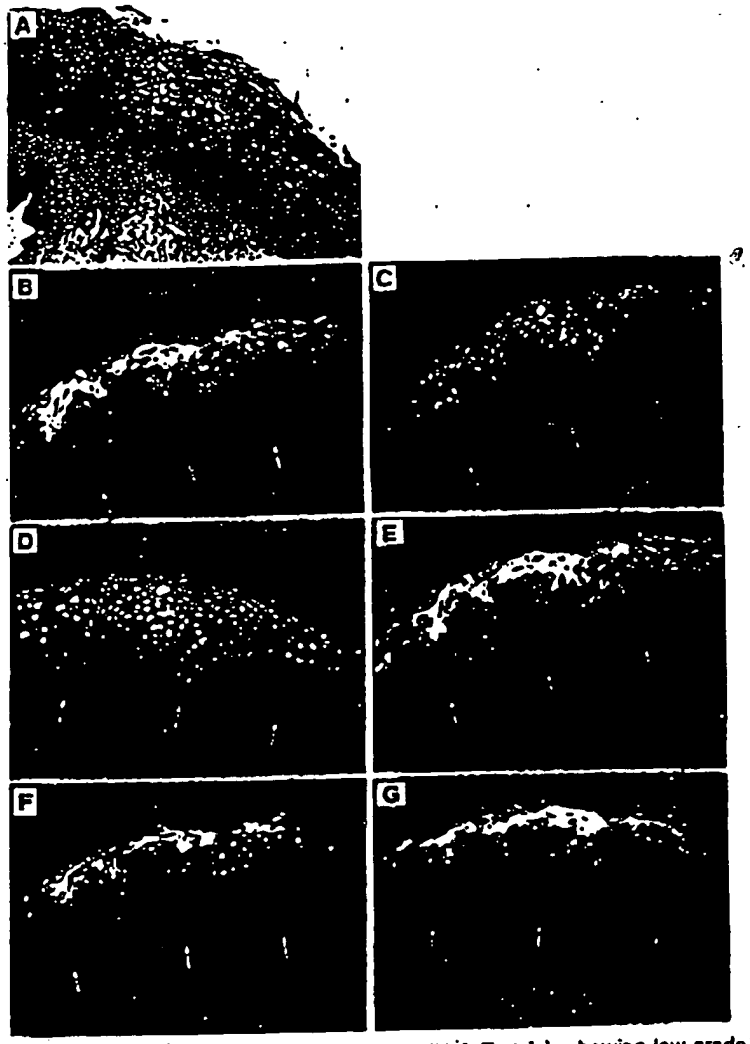

Figure 4 HPV-16 exon-specific probes of CEN I. (*A*) Histopathology (H & E stain), showing low-grade dysplasia and superficial koilocytes, visualized in bright-field. (*B–G*) Dark-field visualization of exposed silver grains in biopsies subjected to in situ hybridization with ^{3}H-labeled antisense RNA probes. Basement membranes are marked with arrows. (*B*) E6-E7; (*C*) E1; (*D*) E2 (note the nuclear localization of signals in *C* and *D*, presumably residual introns and primary transcripts); (*E*) E4-E5; (*F*) L2; and (*G*) L1.

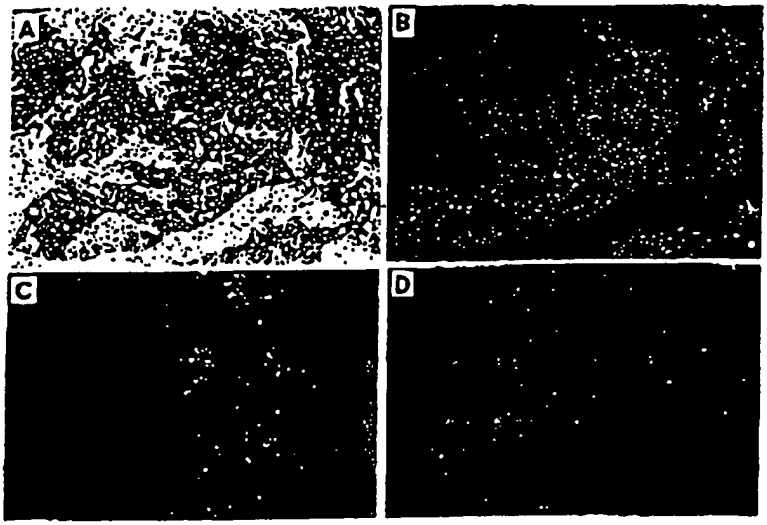

Figure 5 HPV-16 exon-specific probes of invasive adenosquamous carcinoma of the uterine corpus. (*A*) Histology; (*B*) E6-E7; (*C*) E4-E5; and (*D*) L1.

HPV-11 infections. This might suggest that the putative promoter immediately preceding the E1 ORF of HPV-16 is less active than that of HPV-6 and HPV-11. As with HPV-6 and HPV-11, the E1 and the E2 signals were almost exclusively nuclear (Fig. 4C,D). We attribute these to unprocessed primary transcripts and to residual intron material derived from the E6-E7˜E4-E5 or E6-E7˜E2-E5 mRNAs. Cytoplasmic signals from E1 or E2 were low or negligible. Cytoplasmic L2 and L1 ORF signals emerged only in the highly differentiated superficial strata (Fig. 4F,G), consistent with their synthesis as late messages encoding structural proteins. Occasional nuclear L2 and L1 signals in mid-epithelium may represent the 3′ run-on of early transcripts processed at the early polyadenylation site or perhaps are precursors to late messages. The tissue, though typically hyperproliferative, showed full differentiation into stratified epithelium as well as koilocytotic atypia indicative of a HPV cytopathic effect.

The higher grades of lesions exhibited dramatically less viral transcription and also were much more restricted in cellular differentiation. Nonetheless, it is important to recognize that the amounts of viral RNA for the degree of differentiation were of similar or higher abundance compared with RNA concentrations in the basal and parabasal cells of equivalent differentiation in the CIN I. These observations are critical to the proposed mechanisms of HPV carcinogenesis to be de-

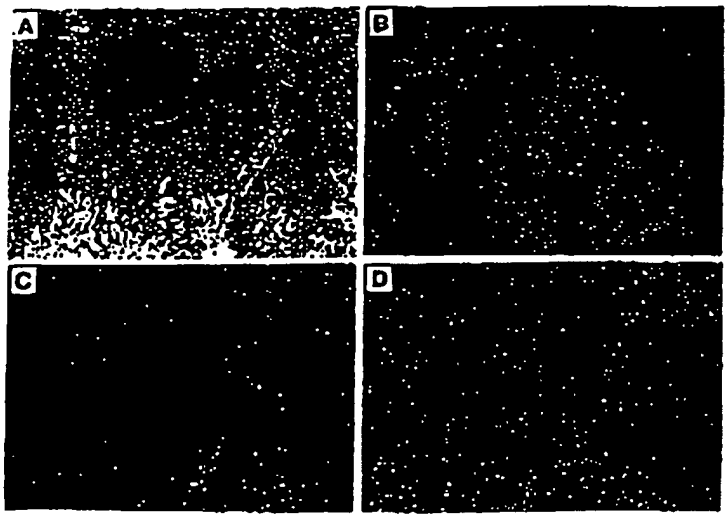

Figure 6 HPV-16 exon-specific probes of invasive squamous carcinoma of the cervix. (*A*) Histology; (*B*) E6-E7; (*C*) E4-E5; and (*D*) L1.

scribed. As might be anticipated, based on a requirement for terminal differentiation, late RNA (L2 and L1) was generally not made in any of the higher-grade lesions. The CIN III exhibited modest amounts of E6-E7 and E4-E5 exons in the cytoplasm and proportionally lower E1 and E2 signals in the nuclei (data not shown). L2 and L1 exons were absent. Similarly, the histologically distinct squamous carcinoma in situ, adenocarcinoma in situ, and invasive adenosquamous carcinoma from a single patient all had the same exon transcription patterns, suggestive of a clonal origin for these tumor types. Within this set of specimens and others we have investigated, E6-E7 signals were clearly more abundant than E4-E5 signals (Fig. 5). One interpretation is that some of the viral DNAs are integrated near the E1/E2 junction, dissociating the E2-E4-E5 region from their promoters, whereas others remain episomal. As described above, we believe that the full-length E2 protein in the less-differentiated cells acts both as an activator and as a repressor for the E6 promoter. Upon deletion or disruption of the E2 and E2-C genes, the E6 promoter would be derepressed, which could account for the higher relative E6-E7 transcription compared with E4-E5 transcription. Consistent with this hypothesis was our finding of a purely integrated pattern of gene expression that exhibits substantial levels of E6-E7-E1 signals but no E2-E4-E5, L2, or L1 signals (Fig. 6 and data not shown) in a second case of invasive cancer.

We have noticed in several similar cases that the L1 exon is anomalously expressed, despite the absence of E4-E5 and L2 signals. We infer that the L1 RNA (Fig. 1B, I) did not arise from its usual promoter, but rather from an upstream host gene promoter, and is part of a transcript that either runs into the integrated HPV DNA or is spliced to the L1 acceptor site. Termination would be expected at the L1 polyadenylation signal. This suggestive evidence for a host-L1 fusion transcript could also provide an explanation for the successful processing of the E6-E7 messages initiated from integrated viral genomes. They, like most eukaryotic messages, probably require splicing for transport to the cytoplasm and polyadenylation for stability. We speculate that the presumptive, interrupted host cell gene downstream from the integrated viral genome provides the 3′ exon(s) and the polyadenylation signal allowing the completion of a E6-E7-host fusion transcript. In brief, it seems as though the HPV DNA integration event leading to carcinogenic progression may require the chance insertion into a class of genes that is expressed in epithelial cells. This proposal is being tested experimentally.

HPV-18-associated lesions have also been examined with analogous HPV-18 exon-specific probes (Fig. 7).

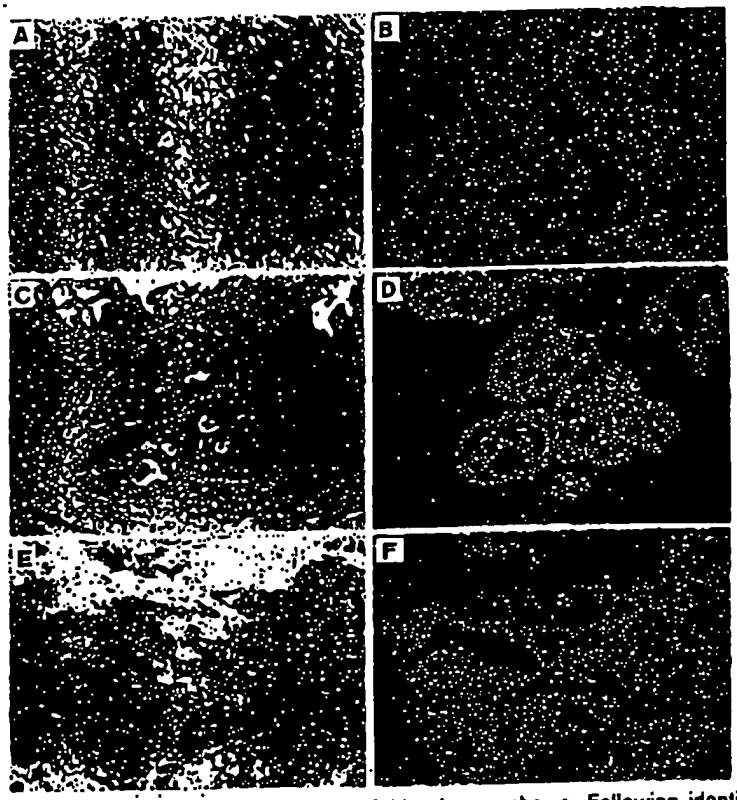

Figure 7 HPV-18-associated carcinomas. Bright-field and dark-field pairs are shown. Following identification with type-specific probes, whole genomic antisense probes were used for greater sensitivity. In each tumor, the total RNA signals are located in invading epithelial cells surrounded by negative stromal fibroblasts. (*A,B*) Squamous carcinoma of the exocervix; (*C,D*) adenocarcinoma of the endocervix; (*E,F*) small-cell neuroendocrine carcinoma of the endocervix.

One case of an HPV-18 CIN I expressed all early and late mRNAs (C. Rhodes et al., unpubl.). HPV-18 has previously been found predominantly in adenocarcinomas and adenosquamous carcinomas (Tase et al. 1988; Wilczynski et al. 1988). We have recently discovered that small cell neuroendocrine carcinomas are highly correlated with the presence and expression of HPV-18 genes (M. Stoler, unpubl.). All carcinomas exhibited a transcription pattern indicative of integration (C. Rhodes et al., unpubl.). The absence of cases with full early region transcription (particularly the E2-E4-E5 region) analogous to some of the HPV-16 CIN III or carcinomas, points to a high probability of rapid HPV-18 DNA integration and loss of the E2-E4-E5 genes. This is entirely consistent with the more aggressive nature of HPV-18 carcinomas relative to those containing HPV-16 DNA (Kurman et al. 1988).

Summary model of HPV carcinogenesis

On the basis of the described transcriptional and regulatory data from experimentation in vitro and in vivo, we propose the following molecular mechanism for HPV carcinogenesis. One of the key elements to the model is that the E2 transcription regulatory protein is translated from a polycistronic transcript that is derived from the E6 promoter and that also encodes several transformation proteins from the E6 and E7 ORFs. Low levels of E2 protein, in association with host transcriptional factors, can up-regulate transcription from the E6 promoter. We hypothesize that the expression of such host factors depends on the state of cellular differentiation. In the mucosotrophic papillomaviruses, a tandem pair of E2 and E2-C protein-binding sites are located adjacent to the E6 promoter TATA motif, which is partially occluded by binding of either E2 protein. Such interference eventually overcomes autostimulation and down-regulates the production of the polycistronic message, providing feedback control to E2 and the viral transforming proteins. This balance leads to a low level of viral maintenance in relatively undifferentiated keratinocytes. As the epithelium differentiates, increases (or other changes) in host regulatory proteins, for instance in factors like CEI and CEII-binding proteins that can overcome E2 repression, may relieve the negative regulation and lead to high levels of early transcription and to viral DNA replication. Increased gene dosage results in higher viral gene expression and the chance for breakthrough to late gene expression and virion production. That cellular transformation does not occur at this stage can be attributed to the fact that these differentiated cells have already lost the ability to divide. In some rare events, the steady maintenance state in the infected stem cells is disrupted. Mutations in the HPV URR or stimulation by other viruses could up-regulate viral early gene expression. Alternatively, mutations in viral genes required for autonomous replication or mutagenic events resulting in viral and host DNA breakage may lead to integration. When integration interrupts the expression of the E2 proteins, the viral transformation proteins would be derepressed, provided the integration happened to occur into active chromatin in a genetic region that could provide downstream RNA splice and polyadenylation sites. In such cases, the dividing basal and parabasal cells would be subject to unusual and excessive amounts of these transforming proteins. Excessive E7 protein stoichiometrically could sequester the RB gene product (Dyson et al., this volume) and perhaps other anti-oncoproteins and release the cells from controlled growth, initiating carcinogensis. Clearly, there are numerous untested aspects of this proposal, but a tantalizing profile of many of the components has been provided through efforts in a number of laboratories studying RNA transcription and processing, enhancer-promoter regulation, protein structure and function, and gene expression in relation to tissue pathogenesis. Without doubt, there are other critical host cell-virus interactions essential for initiation and maintenance of the transformed state and for invasion and metastasis, which will also be the subject of future observation and experimentation.

Acknowledgments

This research was supported by U.S. Public Health Service/National Cancer Institute grants CA-36200 to L.T.C. and T.R.B. and CA-43629 to M.H.S. M.T.C. is a recipient of a U.S. Public Health Service Medical Scientist Training Program Grant (GM-07356) and a March of Dimes Predoctoral Fellowship. S.M.W. was a recipient of a James P. Wilmot Cancer Research Fellowship and an American Cancer Society Institutional Research Award (IN-18). We thank Shirley Thomas for processing the manuscript.

References

Androphy, E.J., D.R. Lowy, and J.T. Schiller. 1987. Bovine papillomavirus E2 *trans*-activating gene product binds to specific sites in papillomavirus DNA. *Nature* 325: 70.

Asselineau, D. and M. Prunieras. 1984. Reconstruction of "simplified" skin: Control of fabrication. *Br. J. Dermatol.* 111: 219.

Baker, C.C., W.C. Phelps, V. Lindgren, M.J. Braun, M.A. Gonda, and P.M. Howley. 1987. Structural and transcriptional analysis of human papillomavirus type 16 sequences in cervical carcinoma cell lines. *J. Virol.* 61: 962.

Banks, L. and L. Crawford. 1988. Analysis of human papillomavirus type 16 polypeptides in transformed primary cells. *Virology* 165: 326.

Broker, T.R. and M. Botchan. 1986. Papillomaviruses: Retrospectives and prospectives. *Cancer Cells* 4: 17.

Broker, T.R. and L.T. Chow. 1986. Human papillomaviruses of the genital mucosa: Electron microscopic analyses of DNA heteroduplexes formed with HPV types 6, 11 and 16. *Cancer Cells* 4: 589.

Chin, M.T., R. Hirochika, H. Hirochika, T.R. Broker, and L.T. Chow. 1988. Regulation of human papillomavirus type 11 enhancer and E6 promoter by activating and repressing proteins from the E2 open reading frame: Functional and biochemical studies. *J. Virol.* 62: 2994.

Choo, K.-B., H.-H. Lee, C.-C. Pan, S.-M. Wu, L.-N. Liew, W.-F. Cheung, and S-H. Han. 1988. Sequence duplication and internal deletion in the integrated human papillomavirus type 16 genome cloned from a cervical carcinoma. *J. Virol.* 62: 1859.

Choo, K.-B., C.-C. Pan, M.S. Liu, H.T. Ng, C.P. Chen, Y.N. Lee, C.F. Chao, C.L. Meng, M.Y. Yeh, and S.H. Han. 1987. Presence of episomal and integrated human papillomavirus DNA sequences in cervical carcinoma. *J. Med. Virol.* 21: 101.

Chow, L.T., M. Nasseri, S.M. Wolinsky, and T.R. Broker. 1987a. Human papillomavirus types 6 and 11 mRNAs from genital condylomata. *J. Virol.* 61: 2581.

Chow, L.T., H. Hirochika, M. Nasseri, M.H. Stoler, S.M. Wolinsky, M.T. Chin, R. Hirochika, D.S. Arvan, and T.R. Broker. 1987b. Human papillomavirus gene expression. *Cancer Cells* 5: 55.

Coppleson, M. and B. Reid. 1967. *Preclinical carcinoma of the cervix uteri.* Pergamon Press, Oxford.

Cripe, T.P., T.H. Haugen, J.P. Turk, F. Tabatabai, P.G. Schmid III, M. Dürst, L. Gissmann, A. Roman, and L.P. Turek. 1987. Transcriptional regulation of the human papillomavirus-16 E6-E7 promoter by a keratinocyte-dependent enhancer, and by viral E2 *trans*-activator and *repressor* gene products: Implications for cervical carcinogenesis. *EMBO J.* 6: 3745.

DeCaprio, J.A., J.W. Ludlow, J. Figge, J.-Y. Shew, C.-M. Huang, W.-H. Lee, E. Marsilio, E. Paucha, and D.M. Livingston. 1988. SV40 large tumor antigen forms a specific complex with the product of the retinoblastoma susceptibility gene. *Cell* 54: 275.

Dürst, M., A. Kleinheinz, M. Hotz, and L. Gissmann. 1985. The physical state of human papillomavirus type 16 DNA in benign and malignant genital tumours. *J. Gen. Virol.* 66: 1515.

Dürst, M., C.M. Croce, L. Gissmann, E. Schwarz, and K. Huebner. 1987. Papillomavirus sequences integrate near cellular oncogenes in some cervical carcinomas. *Proc. Natl. Acad. Sci.* 84: 1070.

Dyson, N., P.M. Howley, K. Munger, and E. Harlow. 1989. The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product. *Science* 243: 934.

El Awady, M.K., J.B. Kaplan, S.J. O'Brien, and R.D. Burk. 1987. Molecular analysis of integrated human papillomavirus 16 sequences in the cervical cancer cell line SiHa. *Virology* 159: 389.

Giri, I. and M. Yaniv. 1988. Study of the E2 gene product of the cottontail rabbit papillomavirus reveals a common mechanism of *trans*activation among papillomaviruses. *J. Virol.* 62: 1573.

Gould, P.R., R.A. Barter, and J.M. Papadimitriou. 1979. An ultrastructural, cytodynamical and autoradiographic study of the mucous membrane of the human cervical canal with reference to subcolumnar cells. *Am. J. Pathol.* 95: 1.

Hirochika, H., T.R. Broker, and L.T. Chow. 1987. Enhancers and *trans*-acting E2 transcriptional factors of papillomaviruses. *J. Virol.* 61: 2599.

Hirochika, H., R. Hirochika, T.R. Broker, and L.T. Chow. 1988. Functional mapping of the human papillomavirus type 11 transcriptional enhancer and its interaction with the *trans*-acting E2 protein. *Genes Dev.* 2: 54.

Hubbert, N.L., J.T. Schiller, D.R. Lowy, and E.J. Androphy. 1988. Bovine papilloma virus-transformed cells contain multiple E2 proteins. *Proc. Natl. Acad. Sci.* 85: 5864.

Kopan, R., G. Traska, and E. Fuchs. 1987. Retinoids as important regulators of terminal differentiation: Examining keratin expression in individual epidermal cells at various stages of keratinization. *J. Cell Biol.* 105: 427.

Koss, L.G. 1979. *Diagnostic cytology and its histopathologic bases.* J.B. Lippincott, Philadelphia.

Kurman, R.J., M.H. Schiffman, W.D. Lancaster, R. Reid, A.B. Jenson, G.F. Temple, and A.T. Lorincz. 1988. Analysis of individual human papillomavirus types in cervical neoplasia: A possible role for type 18 in rapid progression. *Am. J. Obstet. Gynecol.* 159: 293.

Lambert, P.F., B.A. Spalholz, and P.M. Howley. 1987. A transcriptional repressor encoded by BPV-1 shares a common carboxy-terminal domain with the E2 *trans*activator. *Cell* 50: 69.

Land, H., L.F. Parada, and R.A. Weinberg. 1983. Cellular oncogenes and multistep carcinogenesis. *Science* 222: 771.

Lee, W.-H., J.-Y. Shew, F.D. Hong, T.W. Sery, L.A. Donoso, L.-J. Young, R. Bookstein, and E.Y.-H. P. Lee. 1987. The retinoblastoma susceptibility gene encodes a nuclear phosphoprotein associated with DNA binding activity. *Nature* 329: 642.

Lorincz, A.T., W.D. Lancaster, and G.F. Temple. 1986. Cloning and characterization of the DNA of a new human papillomavirus from a woman with dysplasia of the uterine cervix. *J. Virol.* 58: 225.

Matsukura, T., T. Kanda, A. Furuno, H. Yoshikawa, T. Kawana, and K. Yoshiike. 1986. Cloning of monomeric human papillomavirus type 16 DNA integrated within cell DNA from a cervical carcinoma. *J. Virol.* 58: 979.

McBride, A.A., R. Schlegel, and P.M. Howley. 1988. The carboxy-terminal domain shared by the bovine papillomavirus E2 transactivator and repressor proteins contains a specific DNA binding activity. *EMBO J.* 7: 533.

McCance, D.J., R. Kopan, E. Fuchs, and L.A. Laimins. 1988. Human papillomavirus type 16 alters human epithelial cell differentiation in vitro. *Proc. Natl. Acad. Sci.* 85: 7169.

Moll, R., R. Levy, B. Czernobilsky, P. Hohlweg-Majert, G. Dallenbach-Hellweg, and W.W. Franke. 1983. Cytokeratins of normal epithelia and some neoplasms of the female genital tract. *Lab Invest.* 49: 599.

Moskaluk, C. and D. Bastia. 1988. Interaction of the bovine papillomavirus type 1 E2 transcriptional control protein with the viral enhancer: Purification of the DNA-binding domain and analysis of its contact points with DNA. *J. Virol.* 62: 1925.

Nasseri, M., R. Hirochika, T.R. Broker, and L.T. Chow. 1987. A human papilloma virus type 11 transcript encoding an E1^E4 protein. *Virology* 159: 433.

Pater, M.M. and A. Pater. 1985. Human papillomavirus types 16 and 18 sequences in carcinoma cell lines of the cervix. *Virology* 145: 313.

Patten, S.F. 1978. Diagnostic cytopathology of the uterine cervix. *Monogr. Clin. Cytol.* 3: 1.

Pfister, H. 1987. Papillomaviruses: General description, taxonomy, and classification. In *The papillomaviruses* (ed. N.P. Salzman and P.M. Howley), vol. 2, p. 1. Plenum Press, New York.

Phelps, W.C. and P.M. Howley. 1987. Transcriptional *trans*activation by the human papillomavirus type 16 E2 gene product. *J. Virol.* 61: 1630.

Phelps, W.C, C.L. Yee, K. Munger, and P.M. Howley. 1988. The human papillomavirus type 16 E7 gene encodes *trans*activation and transformation functions similar to those of adenovirus E1A. *Cell* 53: 539.

Pirisi, L., S. Yasumoto, M. Feller, J. Doniger, and J.A. DiPaolo. 1987. Transformation of human fibroblasts and keratinocytes with human papillomavirus type 16 DNA. *J. Virol.* 61: 1061.

Riou, G.F., M. Barrois, V. Dutronquay, and G. Orth. 1985. Presence of papillomavirus DNA sequences, amplification of c-*myc* and c-Ha-*ras* oncogenes and enhanced expression of c-*myc* in carcinomas of the uterine cervix. *UCLA Symp. Mol. Cell. Biol. New Ser.* 32: 47.

Sawadogo, M. and R.G. Roeder. 1985. Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. *Cell* 43: 165.

Schneider-Gädicke, A. and E. Schwarz. 1986. Different human cervical carcinoma cell lines show similar transcription patterns of human papillomavirus type 18 early genes. *EMBO J.* 5: 2285.

Schwarz, E., U.K. Freese, L. Gissmann, W. Mayer, B. Roggenbuck, A. Stremlau, and H. zur Hausen. 1985. Structure and transcription of human papillomavirus sequences in cervical carcinoma cells. *Nature* 314: 111.

Seedorf, K., G. Krämmer, M. Dürst, S. Suhai, and W. Röwekamp. 1985. Human papillomavirus type 16 DNA sequence. *Virology* 145: 181.

Shirasawa, H., Y. Tomita, K. Kubota, T. Kasai, S. Sekiya, H. Takamizawa, and B. Simizu. 1986. Detection of human papil-

lomavirus type 16 DNA and evidence for integration into the cell DNA in cervical dysplasia. *J. Gen. Virol.* 67: 2011.

Smotkin, D. and F.O. Wettstein. 1986. Transcription of human papillomavirus type 16 early genes in a cervical cancer and a cancer-derived cell line and identification of the E7 protein. *Proc. Natl. Acad. Sci.* 83: 4680.

———. 1987. The major human papillomavirus protein in cervical cancers is a cytoplasmic phosphoprotein. *J. Virol.* 61: 1686.

Spalholz, B.A., Y.C. Yang, and P.M. Howley. 1985. *Trans*activation of a bovine papilloma virus transcriptional regulatory element by the E2 gene product. *Cell* 42: 183.

Steinberg, B.M. 1986. Laryngeal papillomatosis is associated with a defect in cellular differentiation. *Ciba Found. Symp.* 120: 208.

Stoler, M.H. and T.R. Broker. 1986. In situ hybridization detection of human papilloma virus DNA and messenger RNA in genital condylomas and a cervical carcinoma. *Hum. Pathol.* 17: 1250.

Storey, A., D. Pim, A. Murray, K. Osborn, L. Banks, and L. Crawford. 1988. Comparison of the in vitro transforming activities of human papillomavirus types. *EMBO J.* 7: 1815.

Tase, T., T. Okagaki, B.A. Clark, D.A. Manias, R.S. Ostrow, L.B. Twiggs, and A.J. Faras. 1988. Human papillomavirus types and localization in adenocarcinoma and adenosquamous carcinoma of the uterine cervix: A study by in situ DNA hybridization. *Cancer Res.* 48: 993.

Thierry, F. and M. Yaniv. 1987. The BPV1-E2 *trans*-acting protein can be either an activator or a repressor of the HPV 18 regulatory region. *EMBO J.* 6: 3391.

Whyte, P., K.J. Buchkovich, J.M. Horowitz, S.H. Friend, M. Raybuch, R.A. Weinberg, and E. Harlow. 1988. Association between an oncogene and an anti-oncogene: The adenovirus E1A proteins bind to the retinoblastoma gene product. *Nature* 334: 124.

Wilczynski, S.P., S. Bergen, J. Walker, S.Y. Liao, and L.F. Pearlman. 1988. Human papillomaviruses and cervical cancer: Analysis of histopathologic features associated with different viral types. *Hum. Pathol.* 19: 697.

Woodworth, C.D., P.E. Bowden, J. Doniger, L. Pirisi, W. Barnes, W.D. Lancaster, and J.A. DiPaolo. 1988. Characterization of normal human exocervical epithelial cells immortalized in vitro by papillomavirus types 16 and 18 DNA. *Cancer Res.* 48: 4620.

Yee, C., I. Krishnan-Hewlett, C.C. Baker, R. Schlegel, and P.M. Howley. 1985. Presence and expression of human papillomavirus sequences in human cervical carcinoma cell lines. *Am. J. Pathol.* 119: 361.